

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2001-060956

(43)Date of publication of application : 06.03.2001

(51)Int.Cl.

H04L 12/28

H04L 12/56

H04L 29/10

(21)Application number : 11-235741

(71)Applicant : NIPPON TELEGR & TELEPH CORP
<NTT>

(22)Date of filing : 23.08.1999

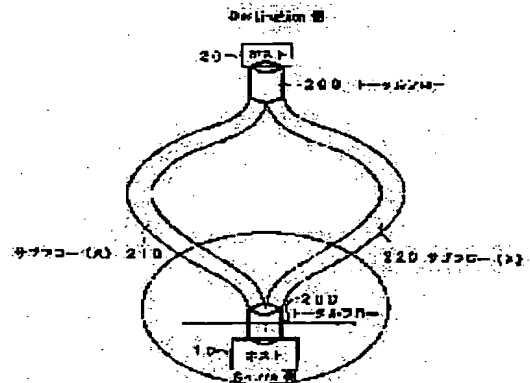
(72)Inventor : HIROTSU TOSHIO
SUGAWARA TOSHIHARU
TAKADA TOSHIHIRO
FUKUDA KENSUKE

(54) TRANSPORT LAYER MULTI-LINK COMMUNICATION METHOD

(57)Abstract:

PROBLEM TO BE SOLVED: To make communication connected with Internet or the like efficient and improve its speed, by satirically supplying a route at every external connection in a network layer, binding each kind of communication executed at every external connection in both end transport layers and supplying them as one kind of communication to a high order layer.

SOLUTION: IP addresses and port numbers are given at the both ends of respective sub-flows (A) 210 and (X) 220 in the same way as TCP. A candidate address and the candidate port number coming them designate an address and a port at the Source side of a total flow 200. Besides, a sequence number concerning a transfer packet is given as the sequence number of TCP. In addition to it, a total sequence number, for example, as the arrangement order of data in the case of binding the sub-flows (A) 210 and (X) 220 is given as the option of TCP. At the side of a Destination, received data is re-constituted in order of the total sequence number.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision
of rejection]

[Date of requesting appeal against examiner's
decision of rejection]

[Date of extinction of right]

*** NOTICES ***

JPO and NCIP are not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. **** shows the word which can not be translated.
3. In the drawings, any words are not translated.

CLAIMS

[Claim(s)]

[Claim 1] The transport layer multi-link correspondence procedure characterized by bundling the communication link for every external connection path by the transport layer, and providing as one communication link to the upper layer in the communication link of the multi-homed connection environment which has two or more external connection to a network.

[Claim 2] The transport layer multi-link correspondence procedure characterized by making the communication link for every external connection path into a subflow, making into a total flow the communication link which bundled this subflow, performing the flow control according to the communication link engine performance of the path concerned by each subflow in a transport layer multi-link correspondence procedure according to claim 1, and performing assignment of the load according to the condition of each subflow, and deletion and an addition of a subflow in a total flow.

[Claim 3] The transport layer multi-link correspondence procedure characterized by giving the total sequence number which shows the sequence at the time of bundling the communication link for every path to a transfer packet in addition to the sequence number of the communication link for every path in claim 1 and the transport layer multi-link correspondence procedure of two publications.

[Claim 4] The transport layer multi-link correspondence procedure characterized by increasing every one connection in a transport layer multi-link correspondence procedure according to claim 1 to 3 in case connection of a multi-link is started.

[Claim 5] The transport layer multi-link correspondence procedure characterized by stretching two or more connection in parallel, and bundling it later in a transport layer multi-link correspondence procedure according to claim 1 to 3 in case connection of a multi-link is started.

[Claim 6] The transport layer multi-link correspondence procedure characterized by preparing a physical data link for every external connection, and connecting a host and a router to each data link in a transport layer multi-link correspondence procedure according to claim 1 to 5.

[Claim 7] The transport layer multi-link correspondence procedure characterized by building as a network which divides a single physical interface into two or more imagination logic data links, and leads to each external connection for every logic data link in a transport layer multi-link correspondence procedure according to claim 1 to 5.

[Claim 8] The host and router which are connected through a network in a transport layer multi-link correspondence procedure according to claim 1 to 5 are a transport layer multi-link correspondence procedure characterized by having two or more IP addresses physically or logically, and matching each IP address with each external connection.

[Translation done.]

*** NOTICES ***

JPO and NCIPJ are not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. **** shows the word which can not be translated.
3. In the drawings, any words are not translated.

DETAILED DESCRIPTION

[Detailed Description of the Invention]

[0001]

[Field of the Invention] This invention relates to the communication technology between hosts, routers, etc. which were connected through the Internet or a private network, and relates to the suitable transport layer multi-link correspondence procedure for improvement in communication link effectiveness in detail in the environment of the multi-homed connection which has two or more external connection to the Internet etc.

[0002]

[Description of the Prior Art] Current and a TCP/IP protocol group are used for the communication link between devices, such as a host and a router, in the communication link of the Internet environment. By the TCP/IP protocol group, each protocol is constituted hierarchical, and the configuration consists of a "link layer", "a network layer (IP)", "the transport layer (TCP and UDP)", and the "application layer", as shown in drawing 2. A "link layer" controls a physical side face with a physical interface etc. and a driver. A "network layer" controls migration of the packet from the edge on the Internet to an edge (End-to-End), and IP (Internet Protocol), ICMP (Internet Control Message Protocol), etc. are used. The "transport layer" controls the data flow for two sets of hosts, and multiplexing, and UDP (User Datagram Protocol) which offers TCP (Transmission Control Protocol) which offers a reliable communication link, and a simple datagram transfer is used. The "application layer" performs processing depending on each application.

[0003] Two topologies, the single topology which connects with a single external organization and, on the other hand, performs all Internet communication links through the single external organization in the connection during the organization of the Internet environment, and the multi-homed topology which carries out direct continuation to two or more external organizations, uses those external connection properly suitably, and communicates with the Internet, are used.

[0004] Path control of packet delivery is performed in the Internet by IP which is the protocol of a "network layer." In IP, a path is used alternatively, and the path of the packet which goes to a certain destination is changed as much as possible, is twisted (it is stable), and is usually employed like. In delivery of the packet addressed to the network which is making multi-homed connection in a multi-homed connection environment, the network which is making multi-homed connection publicizes the path information addressed to a self-network to each of two or more external connection, it is that an external network determines a delivery path based on the path information, and selection of two or more paths is performed. Moreover, in the delivery of the packet from a network which is making multi-homed connection, based on the path information acquired from each of two or more external connection, it is determining a path in a self-network, and selection from two or more paths is performed.

[0005] In the data link layer, it considers as the technique accelerated using a low speed link two or more, and the multi-link technique is used. It realizes as an escape of PPP (Point-to-Point Protocol), and this is used for governing two or more mainly homogeneous data links.

[0006]

[Problem(s) to be Solved by the Invention] although a path is chosen by the network layer according to the framework of the present "path control in a network layer" in case it communicates with the specific partner of a network to the external network which is making multi-homed connection — this — the inside of a network — being dynamic (time) — since it cannot respond to the situation of the traffic to change, the path is not necessarily the optimal path to the communication link of the transport layer. That is, the path control of a network layer is the path control according to the weight, priority, path length, etc. of pass, and the path which these do not necessarily express the condition of having taken the actual congestion of the moment etc. into consideration, and was chosen is not not necessarily the maximum high speed.

[0007] Moreover, if it sees from a multi-homed network, in order to use alternatively from two or more paths, two or more external connection cannot be used up effectively. Therefore, the path cannot be used even if allowances are in bands other than the path currently used for the present communication link. Furthermore, routing information increases as compared with the number of the networks which make multi-homed connection, and a burden is placed on routing.

[0008] The purpose of this invention is in a multi-homed connection environment to enable the selection of a dynamic communication path and the efficient use of two or more external connection according to the situation of actual data transfer, and realize increase in efficiency of the communication link between hosts, routers, etc. which were connected through the Internet or a private network, and improvement in the speed.

[0009]

[Means for Solving the Problem] In a network layer, this invention offers statically the path for two or more external connection of every, and offers in a bundle each communication link (each link) which carried out for every external connection path as one communication link to (this is called multi-link) and the upper layer by the transport layer of both ends.

[0010] The flow control according to the communication link engine performance of a path is performed by the communication link (subflow) for every path, and deletion and an addition of a subflow are processed distribution of the burden according to the condition of each subflow, and if needed by the total flow which bundled each subflow.

[0011] This invention enables it to communicate using the optimal path according to the communication link effectiveness at the time of the communication link by the transport layer out of two or more paths in a multi-homed connection environment. Moreover, it becomes possible to communicate efficiently using two or more paths.

[0012]

[Embodiment of the Invention] Hereafter, a drawing explains the gestalt of 1 operation of this invention. Drawing 1 is the conceptual diagram of the protocol stack in the transport layer multi-link communication link of this invention. Here, the transport layer is divided into TCP, UDP, and a multi-link although the configuration of a ring layer, a network layer, the transport layer, and the application layer is fundamentally [as drawing 2] the same. TCP of a link layer, a network layer, and the transport layer and the group of UDP exist virtually for every path of two or more external connection (drawing 1 3 sets), bundle each communication link (link) performed for every path in the multi-link layer of the transport layer (multi-link), and offer it as one communication link to the application layer of a high order. Below, the example which extends TCP and realizes the transport layer multi-link communication link of this invention is explained in full detail.

[0013] Each host using a transport layer multi-link communication link has two or more IP addresses physically or logically. Two or more of these IP addresses that one host has are equivalent to each external connection, respectively, and serve as external connection to which the path used for a packet transfer according to a sending agency IP address corresponds. That is, two or more paths in the combination of two or more sending agency IP addresses (source address) and a destination IP address (destination address) are realized.

[0014] The conceptual diagram of a transport layer multi-link communication link in case external connection is two as an example is shown in drawing 3 . In drawing 3 , 10 takes as the host of an origination side (source side), and 20 is taken as the host of a receiving side

(destination side). The host 10 is connected to the external networks A and X through the router 101 and the router 102, respectively. Corresponding to this the external connection of each, the host 10 has two IP addresses, X. "A. B.s1.s2" and "Y.t1.t2." A host's 20 IP address is set to "D." A packet 110 is transmitted to a host 20 in the path of a router 101 and Network A from a host 10 with the combination of the transmitting agency address (source address) "A. B.s1.s2" and a destination address (destination address) "D." Moreover, a packet 120 is transmitted to a host 20 in the path of a router 102 and Network X from a host 10 with the combination of the transmitting agency address "X. Y.t1.t2" and a destination address "D." In addition, although a transmitting agency port number (source port number) and a destination port number (destination port number) also exist in a TCP/IP packet, in drawing 3, it has omitted for simplification.

[0015] Here, the communication link of the end to end which bundled a subflow, and a call and its subflow for the communication link for every path which passes along each external connection is called a total flow. Application uses this total flow like the conventional communication link.

[0016] The conceptual diagram of the relation between a subflow and a total flow is shown in drawing 4. In drawing 4, 200 is a total flow and 210 and 220 are the subflows A and X. Each subflows 210 and 220 perform a usual flow control and a usual congestion control corresponding to the communication link engine performance of the path concerned like TCP, and the total flow 200 performs processing called assignment to the transfer data to a subflow, and deletion and an addition of a subflow according to the information on the communication band of each subflows 210 and 220 etc.

[0017] Next, the processing in TCP is explained concretely. The same usual IP address and same usual port number as TCP are attached to the both ends (the source, destination) of each subflow. Here, one IP address in the IP address attached to two or more subflows for every communicative edge is called the representation address. Moreover, a port number calls any one a representation port number similarly. This representation address and a representation port number are used as what specifies the address port by the side of the local of a total flow (source side). Moreover, as a sequence number of TCP, the sequence number (subflow sequence number) about the transfer packet for every subflow is given, and each subflow is processed as 1 connection of the conventional TCP. In addition, the data at the time of bundling a subflow arrange, and a total sequence number is given as an option of TCP as order. This total sequence number expresses the sequence of data before dividing into each subflow, and reconfigures the data received by each subflow in order of this total sequence number in a receiving side.

[0018] The example of the factice / total sequence number given to drawing 5 at the transfer packet of each subflow is shown. Data 1, data 2, and data 5 being the paths of the subflow (A) 210, and data 3 and data 4 being the paths of the subflow (X) 220, and transmitting the data of a configuration as this shows drawing 6 to a host 10 to the host 20, respectively is shown. In this case, in the host 10 by the side of the source, 1, 2, and 3 are attached to data 1, data 2, and data 5 as a sequence number (subflow sequence number) of TCP, respectively, and, similarly 1 and 2 are attached to data 3 and data 4 as a sequence number of TCP, respectively. This is the same as usual. In addition, in a host 10, the total sequence numbers 1-5 are attached to data 1 - data 5 using the option of TCP. Although omitted in drawing 5, the group of the IP address port number of the source/destination is given to each packet, thereby, data 1, data 2, and data 5 are transmitted in the path of the subflow (A) 210, and data 3 and data 4 are transmitted in the path of the subflow (X) 220. In the host 20 by the side of a destination, the data 1-5 which show the data received by each SABUFURO 210,220 course to drawing 6 according to the total sequence number to which it is given by each data in the bundle as a total flow 200 are reconfigured.

[0019] In addition, in drawing 5, the group "A. B.s1.s2, P1" of the IP address enclosed with a broken line and a port number shows the representation address and a representation port number. Of course, it is possible to make X. other "Y.t1.t2, P2" into the representation address and a representation port number.

[0020] Connection of each subflow is the same as that of handshaking of TCP. Two or more subflows are started, it is sufficient in a bundle, and there are five, SETUP, ATTACH, DETACH, SHUTDOWN, and RECOVERY, as a procedure of releasing the subflow which became

unnecessary.

[0021] SETUP is the processing which starts connection of a multi-link, and has sequential setup which increases every one connection of a subflow, and concurrent setup which stretches connection of two or more subflows in parallel, and bundles it later.

[0022] In Sequential Setup, the option of MP(Multi-link Extension Protocol)_REQUERST is first attached about the subflow which passes along a one-eyed path, and the setup by 3-way handshake of TCP is performed. When the communications partner supports the transport layer multi-link, ACK which attached the option of MP_GLANTED on the occasion of 3-way handshake is returned. The identifier (flow identifier) which identifies a total flow is given to the message of this MP_GRANTED (every one ID per each direction of connection is given). When connection with a one-eyed path goes wrong, the subflow which passes along other paths is chosen, and the same setup is performed. In Sequential Setup, the local address port of connection of this beginning turns into the representation address and a representation port. If the set of a subflow is completed, a communication link will be started using the subflow (it adds to the usual sequence number at each packet, and a total sequence number is *****). About the subflow of other paths, the subflow is added by ATTACH processing described later. The conceptual diagram of Sequential Setup is shown in drawing 7.

[0023] On the other hand, in Concurrent Setup, connection of the subflow of the number of larger arbitration than one of the paths which have more than one is set up in coincidence. In this case, the address and the port number by the side of a local are first acquired about each path set up in coincidence. And a flow identifier is attached to the option of MP_REQUEST in the case of the setup of each subflow. This flow identifier is used [that two or more TCP connection (two or more subflows) is one multilinked connection and] for a discernment sake. When the direction which receives connection supports the transport layer multi-link, the other party address and the port of a connection request which arrived first are made into the representation address and the representation port of the other party (connection initiation side) of the connection, the representation address and representation port, and flow identifier are attached to the option of MP_GRANTED, and ACK of 3-way handshake is returned. About the subflow after the second, MP_GRANTED which attached the representation address and the representation port of a total flow, and the flow discernment ID is returned with ACK. The conceptual diagram of Concurrent Setup is shown in drawing 8.

[0024] A subflow is added in ATTACH (refer to drawing 7). In the case of ATTACH, the option and flow identifier of MP_ATTACH are attached to the partner corresponding to a multi-link, and a subflow is connected at it. If the total flow expressed in the representation address and a representation port is in a still open condition, the other party will attach MP_ATTACHED and will return ACK. If the setup of a subflow is completed, some data of a total flow will be started (in addition to the usual sequence number, a total sequence number is henceforth attached to each packet). The addition of the subflow by this ATTACH can consider generating in the time of changing the condition of other subflows sharply, the periodic trigger in fixed time amount, etc., when the amount of transfers of a total flow increases suddenly besides above-mentioned Sequential Setup.

[0025] DETACH deletes a subflow. Deletion of the subflow by DETACH is the same as termination of connection of the usual TCP about each subflow. However, the port number is not reusable until a total flow is completed, in case the subflow currently used as a representation port number is DETACH(ed). As compared with other flows, the case where the total engine performance does not change even if it adds a subflow, and a transfer generate deletion of a subflow, when small.

[0026] SHUTDOWN is processing which completes the communication link by the total flow. In the case of SHUTDOWN, while ending each subflow like the usual TCP, when the port number concerned is being used to already have ended the subflow as a representation port number, the port number is released.

[0027] RECOVERY is processing performed when a subflow is cut suddenly. Although what is necessary is just to arrange data in order of a total sequence number in a total flow since each subflow is performing resending processing in the steady state, only when a subflow is suddenly

cut in the middle of a communication link, resending processing is carried out using the data after the time of having arrived not being checked as RECOVERY processing as a total flow.

[0028] More nearly finally than the network layer by which the need is carried out, in order to realize the above processing, a lower layer is explained.

[0029] In order to realize the transport layer multi-link communication link of this invention which was stated above, it is necessary to constitute the network of two or more network layers for every external connection. For this solution, means, such as "multiplexing of a physical data link", "multiplexing of a logic data link", and "multiplexing of IP network", can be considered. As shown in drawing 9, "multiplexing of a physical data link" prepares every one physical data link for every external connection, and a host connects it to each data link. [two or more] For example, it is the gestalt which one host 10 has two or more Ethernet (trademark) interfaces, and connects each to a different physical network. As shown in drawing 10, using the virtual LAN (Virtual LAN, VLAN) function which a data link layer prepares, "multiplexing of a logic data link" divides a single physical interface into two or more imagination logic data links, and builds it as a network which leads to each external connection for every logic data link. Specifically, it is Ethernet (trademark) looked at by IEEE802.1Q. VC (VirtualCircuit) of VLAN or ATM is equivalent to this. As shown in drawing 11, "multiplexing of IP network" gives each host two or more IP addresses for every external connection, and a middle router carries out path control with *** of the IP address of a sending agency, and the IP address of a transmission place. Specifically, froute, SuMiRe, etc. are mentioned as such router ability.

[0030] As mentioned above, although this invention was concretely explained based on the operation gestalt, as for this invention, it is needless to say for it to be able to change variously in the range which is not limited to this operation gestalt and does not deviate from the summary. For example, when the router has two or more addresses physically or logically, a multi-link communication link is possible, although the communication link between hosts was assumed with the operation gestalt similarly between routers. Moreover, although the operation gestalt explained the example which extends TCP and realizes multilinked connection, even if it extends UDP, multilinked connection is realizable similarly.

[0031]

[Effect of the Invention] As mentioned above, according to this invention, in the communication link to an external network with two or more external connection from a network, efficient use of two or more external connection and selection of the dynamic communication path according to the situation of actual data transfer are attained by bundling the communication link for two or more external connection paths of every by the transport layer. Thereby, the communication link in the Internet etc. can be more efficiently performed at a high speed again.

[Translation done.]

*** NOTICES ***

JPO and NCIPi are not responsible for any damages caused by the use of this translation.

- 1.This document has been translated by computer. So the translation may not reflect the original precisely.
- 2.**** shows the word which can not be translated.
- 3.In the drawings, any words are not translated.

DESCRIPTION OF DRAWINGS

[Brief Description of the Drawings]

[Drawing 1] It is the conceptual diagram of the protocol stack of a transport layer multi-link communication link of this invention.

[Drawing 2] It is drawing showing the protocol stack in the usual TCP/IP sweet.

[Drawing 3] It is drawing showing the address of a transport layer multi-link communication link of this invention, and the relation of a packet transfer.

[Drawing 4] It is drawing showing the relation between the subflow of a transport layer multi-link communication link of this invention, and a total flow.

[Drawing 5] It is drawing showing the relation between the subflow sequence number of a transport layer multi-link communication link of this invention, and a total flow sequence number.

[Drawing 6] It is drawing showing the data configuration for explaining the relation between the subflow sequence number of drawing 5 , and a total sequence number.

[Drawing 7] Sequential in the transport layer multi-link communication link of this invention It is the conceptual diagram of Setup.

[Drawing 8] Concurrent in the transport layer multi-link communication link of this invention It is the conceptual diagram of Setup.

[Drawing 9] It is drawing showing the multiplexing configuration of the physical data link which realizes the transport layer multi-link communication link of this invention.

[Drawing 10] It is drawing showing the multiplexing configuration of the logic data link which realizes the transport layer multi-link communication link of this invention.

[Drawing 11] It is drawing showing the multiplexing configuration of IP network which realizes the transport layer multi-link communication link of this invention.

[Description of Notations]

10 20 Host

30 Network

101,102 Router

110,120 Packet

200 Total Flow

210,220 Subflow

[Translation done.]

(51) IntCl. ⁷	識別記号	F I	テーマコード(参考)
H 0 4 L	12/28	H 0 4 L 11/00	3 1 0 Z 5 K 0 3 0
	12/56	11/20	1 0 2 A 5 K 0 3 3
	29/10	13/00	3 0 9 Z 5 K 0 3 4
			9 A 0 0 1

審査請求 未請求 請求項の数 8 O L (全 8 頁)

(21) 出願番号 特願平11-235741

(22) 出願日 平成11年8月23日 (1999.8.23)

(71) 出願人 000004226

日本電信電話株式会社

東京都千代田区大手町二丁目3番1号

(72) 発明者 廣津 登志夫

東京都千代田区大手町二丁目3番1号 日

本電信電話株式会社内

(72) 発明者 菅原 俊治

東京都千代田区大手町二丁目3番1号 日

本電信電話株式会社内

(74) 代理人 100073760

弁理士 鈴木 誠

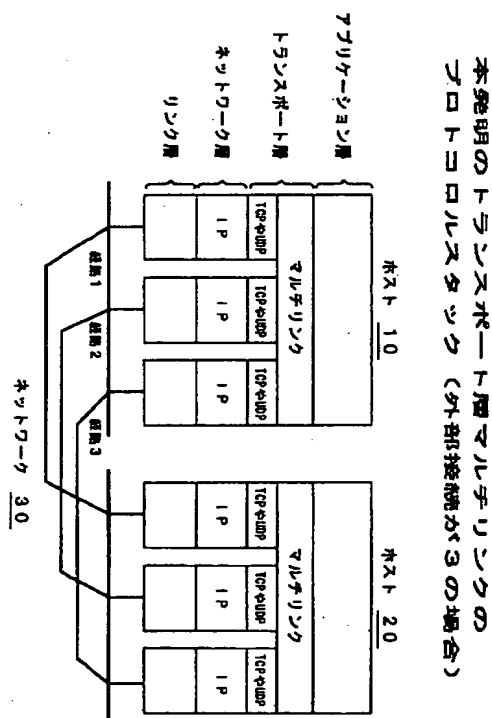
最終頁に続く

(54) 【発明の名称】 トランスポート層マルチリンク通信方法

(57) 【要約】

【課題】 マルチホーム接続の環境において、複数の外部接続の効率的利用、実際のデータ転送状況に応じた動的通信経路の選択を可能にする。

【解決手段】 ネットワーク層では複数の外部接続毎の経路を静的に提供し、各外部接続経路毎に行われる通信を両端のトランスポート層で束ね（マルチリンク）、アプリケーション層に対しては一つの通信として提供する。経路の通信性能に応じたフロー制御は、各経路毎の通信（サブフロー）で通常のTCPと同様にして処理し、各サブフローを束ねたトータルフローにおいて、各サブフローの状態に応じた負担の分散や、サブフローの削除・追加を処理する。



【特許請求の範囲】

【請求項1】 ネットワークに対して複数の外部接続を持つマルチホーム接続環境の通信において、各々の外部接続経路毎の通信をトランスポート層で束ねて、上位層に対して一つの通信として提供することを特徴とするトランスポート層マルチリンク通信方法。

【請求項2】 請求項1記載のトランスポート層マルチリンク通信方法において、各々の外部接続経路毎の通信をサブフローとし、該サブフローを束ねた通信をトータルフローとし、各サブフローでは当該経路の通信性能に応じたフロー制御を行い、トータルフローでは各サブフローの状態に応じた負荷の割り振りやサブフローの削除・追加を行うことを特徴とするトランスポート層マルチリンク通信方法。

【請求項3】 請求項1、2記載のトランスポート層マルチリンク通信方法において、転送パケットに、各経路毎の通信のシーケンス番号に加えて、各経路毎の通信を束ねる際の順番を示すトータルシーケンス番号を付与することを特徴とするトランスポート層マルチリンク通信方法。

【請求項4】 請求項1乃至3記載のトランスポート層マルチリンク通信方法において、マルチリンクの接続を立ち上げる際に、接続を一つずつ増やしていくことを特徴とするトランスポート層マルチリンク通信方法。

【請求項5】 請求項1乃至3記載のトランスポート層マルチリンク通信方法において、マルチリンクの接続を立ち上げる際に、複数の接続を並行に張って後で束ねることを特徴とするトランスポート層マルチリンク通信方法。

【請求項6】 請求項1乃至5記載のトランスポート層マルチリンク通信方法において、各々の外部接続毎に物理データリンクを用意し、エンドホストやルータをそれぞれのデータリンクに接続することを特徴とするトランスポート層マルチリンク通信方法。

【請求項7】 請求項1乃至5記載のトランスポート層マルチリンク通信方法において、単一の物理インタフェースを複数の仮想的な論理データリンクに分割し、各々の論理データリンク毎に各外部接続につながるネットワークとして構築することを特徴とするトランスポート層マルチリンク通信方法。

【請求項8】 請求項1乃至5記載のトランスポート層マルチリンク通信方法において、ネットワークを介して接続されるホストやルータは物理的もしくは論理的に複数のIPアドレスを持ち、各IPアドレスを各外部接続に対応づけることを特徴とするトランスポート層マルチリンク通信方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は、インターネットやプライベートネットワークを介して接続されたホスト、

ルータ等の間での通信技術に係り、詳しくは、インターネット等に対して複数の外部接続を持つマルチホーム接続の環境において、通信効率の向上に好適なトランスポート層マルチリンク通信方法に関する。

【0002】

【従来の技術】 インターネット環境の通信においては、現在、TCP/IPプロトコル群がホストやルータといった機器の間での通信に使われている。TCP/IPプロトコル群では、各プロトコルは階層的に構成されており、その構成は、図2に示すように、「リンク層」「ネットワーク層(IP)」「トランスポート層(TCPやUDP)」「アプリケーション層」から成っている。

「リンク層」は物理的なインタフェース等やドライバとの物理的側面を制御する。「ネットワーク層」はインターネット上の端から端(End-to-End)までのパケットの移動を制御し、IP(Internet Protocol)やICMP(Internet Control Message Protocol)等が使われている。「トランスポート層」は二台のホスト間のデータの流れや多重化を制御し、信頼性の高い通信を提供するTCP(Transmission Control Protocol)や単純なデータグラム転送を提供するUDP(User Datagram Protocol)が使われている。「アプリケーション層」は各アプリケーションに依存した処理を行う。

【0003】 一方、インターネット環境の組織間の接続においては、単一の外部組織と接続して全てのインターネット通信をその単一の外部組織を通じて行う単一接続形態と、複数の外部組織と直接接続して適宜それらの外部接続を使い分けてインターネットと通信するマルチホーム接続形態の二つの接続形態が使われている。

【0004】 インターネットでは、「ネットワーク層」のプロトコルであるIPにより、パケット配送の経路制御が行われている。IPでは、経路が選択的に利用され、通常、ある目的地に向かうパケットの経路はできるだけ変動しない(安定な)ように運用されている。マルチホーム接続環境でのマルチホーム接続しているネットワーク宛のパケットの配送においては、マルチホーム接続しているネットワークが複数の外部接続のそれぞれに自ネットワーク宛の経路情報を広報し、外部ネットワークがその経路情報をもとに配送経路を決定することで、複数の経路の選択が行われている。また、マルチホーム接続しているネットワークからのパケットの配送においては、複数の外部接続のそれぞれから獲得する経路情報をもとに、自ネットワークで経路を決定することで、複数の経路からの選択が行われている。

【0005】 データリンク層に於ては、低速なリンクを複数利用して高速化する手法としてマルチリンク技術が利用されている。これは、PPP(Point-to-Point Protocol)の拡張として実現されており、主に均質な複数のデータリンクを束ねるのに利用されている。

【0006】

【発明が解決しようとする課題】現状の「ネットワーク層における経路制御」のフレームワークでは、マルチホーム接続をしているネットワークから外部ネットワークの特定の相手と通信する際に、ネットワーク層で経路を選択するが、これはネットワーク中の動的（時間的）に変動するトラフィックの状況に対応できないために、その経路がトランスポート層の通信に対して最適な経路とは限らない。すなわち、ネットワーク層の経路制御は、パスのウエイト・優先度・パス長等に従った経路制御であり、これらはその瞬間の実際の輻輳等を考慮した状態を表わしている訳ではなく、選択された経路が必ずしも最高速とは限らない。

【0007】また、マルチホームネットワークから見ると、複数の経路から選択的に利用するために、複数の外部接続を有効に使い切ることができない。そのため、現在の通信に使っている経路以外の帯域に余裕があっても、その経路を利用することができない。さらに、マルチホーム接続するネットワークの数に比してルーティング情報が増大し、ルーティングに負担がかかる。

【0008】本発明の目的は、マルチホーム接続環境において、実際のデータ転送の状況に応じた動的な通信経路の選択や複数の外部接続の効率的利用を可能にして、インターネットやプライベートネットワークを介して接続されたホスト、ルータ等の間の通信の効率化、高速化を実現することにある。

【0009】

【課題を解決するための手段】本発明は、ネットワーク層では複数の外部接続毎の経路を静的に提供し、各外部接続経路毎に行った各々の通信（各リンク）を両端のトランスポート層で束ね（これをマルチリンクと言う）、上位層に対しては一つの通信として提供する。

【0010】経路の通信性能に応じたフロー制御は、各経路毎の通信（サブフロー）で行い、各サブフローを束ねたトータルフローで各サブフローの状態に応じた負担の分散や、必要に応じてサブフローの削除・追加を処理する。

【0011】本発明により、マルチホーム接続環境において、複数の経路の中から、トランスポート層での通信時の通信効率に従って最適な経路を利用して通信することが可能になる。また、複数の経路を利用して効率良く通信することが可能になる。

【0012】

【発明の実施の形態】以下、本発明の一実施の形態を図面により説明する。図1は、本発明のトランスポート層マルチリンク通信におけるプロトコルスタックの概念図である。ここで、リンク層、ネットワーク層、トランスポート層、アプリケーション層の構成は図2と基本的に同様であるが、トランスポート層がTCPやUDPとマルチリンクとに分かれる。リンク層、ネットワーク層及びトランスポート層のTCPやUDPの組は、複数の外

部接続の各経路毎に仮想的に存在し（図1では3組）、各経路毎に行われる各々の通信（リンク）をトランスポート層のマルチリンク層で束ねて（マルチリンク）、上位のアプリケーション層に対しては一つの通信として提供する。以下では、TCPを拡張して本発明のトランスポート層マルチリンク通信を実現する実施例について詳述する。

【0013】トランスポート層マルチリンク通信を利用する各ホストは、物理的もしくは論理的に複数のIPアドレスを持っている。一つのホストが持つ、この複数のIPアドレスは、それぞれ各外部接続に対応しており、発信元IPアドレスに応じてパケット転送に使用される経路が対応する外部接続となる。すなわち、複数の発信元IPアドレス（ソースアドレス）と宛先IPアドレス（デスティネーションアドレス）の組合せで複数の経路が実現する。

【0014】図3に、一例として外部接続が2つの場合のトランスポート層マルチリンク通信の概念図を示す。図3において、10は発信側（ソース側）のホスト、20は受信側（デスティネーション側）のホストとする。ホスト10はそれぞれルータ101とルータ102を通して外部ネットワークA、Xに接続されている。このそれぞれの外部接続に対応して、ホスト10は「A. B. s1. s2」と「X. Y. t1. t2」の2つのIPアドレスを持っている。ホスト20のIPアドレスは「D」とする。パケット110は、送信元アドレス（ソースアドレス）「A. B. s1. s2」と宛先アドレス（デスティネーションアドレス）「D」の組合せにより、ホスト10からルータ101、ネットワークAの経路でホスト20に転送される。また、パケット120は、送信元アドレス「X. Y. t1. t2」と宛先アドレス「D」の組み合わせにより、ホスト10からルータ102、ネットワークXの経路でホスト20に転送される。なお、TCP/IPパケットには、送信元ポート番号（ソースポート番号）と宛先ポート番号（デスティネーションポート番号）も存在するが、図3では簡単化のために省略してある。

【0015】ここで、それぞれの外部接続を通る経路毎の通信をサブフローと呼び、そのサブフローを束ねたエンド・ツー・エンドの通信をトータルフローと呼ぶ。アプリケーションは、このトータルフローを従来の通信と同様に利用する。

【0016】図4に、サブフローとトータルフローの関係の概念図を示す。図4において、200がトータルフロー、210と220がサブフローA、Xである。各サブフロー210、220は通常のTCPと同様に当該経路の通信性能に応じたフロー制御や輻輳制御を行い、トータルフロー200は各サブフロー210、220の通信帯域等の情報に応じて、サブフローへの転送データへの割り振りやサブフローの削除・追加という処理を行

う。

【0017】次に、TCPにおける処理を具体的に説明する。各サブフローの両端（ソース、デスティネーション）には、通常のTCPと同様のIPアドレスとポート番号が付いている。ここで、通信の各端毎に複数のサブフローに付いたIPアドレスの中の一つのIPアドレスを代表アドレスと呼ぶ。また、ポート番号も同様にいずれか一つを代表ポート番号と呼ぶ。この代表アドレスと代表ポート番号はトータルフローのローカル側（ソース側）のアドレス・ポートを指定するものとして利用する。また、TCPのシーケンス番号としては、各サブフロー毎の転送パケットに関するシーケンス番号（サブフローシーケンス番号）が付与され、各サブフローが従来のTCPの一接続として処理される。これに加えて、サブフローを束ねる際のデータの並べ順として、トータルシーケンス番号をTCPのオプションとして付与する。このトータルシーケンス番号は、各サブフローに分割する前のデータの順番を表すもので、受信側では各サブフローで受け取ったデータを、このトータルシーケンス番号の順に再構成する。

【0018】図5に、各サブフローの転送パケットに付与されるサブ/トータルシーケンス番号の具体例を示す。これは、図6に示すような構成のデータをホスト10からホスト20に対して、データ1とデータ2とデータ5はサブフロー（A）210の経路で、データ3とデータ4はサブフロー（X）220の経路で、それぞれ転送することを示している。この場合、ソース側のホスト10では、TCPのシーケンス番号（サブフローシーケンス番号）として1、2、3をデータ1、データ2及びデータ5にそれぞれ付け、同じくTCPのシーケンス番号として1、2をデータ3、データ4にそれぞれ付ける。これは従来と同様である。これに加えて、ホスト10では、データ1～データ5に、トータルシーケンス番号1～5をTCPのオプションを利用して付ける。図5では省略したが、各パケットにはソース/デスティネーションのIPアドレス・ポート番号の組が付され、これにより、データ1とデータ2とデータ5はサブフロー（A）210の経路で転送され、データ3とデータ4はサブフロー（X）220の経路で転送される。デスティネーション側のホスト20では、各サブフロー210、220経由で受け取ったデータをトータルフロー200として束ね、各データに付与されているトータルシーケンス番号に従って図6に示すデータ1～5を再構成する。

【0019】なお、図5において、破線で囲ったIPアドレスとポート番号の組「A. B. s1. s2, P1」は代表アドレス・代表ポート番号を示している。勿論、他の「X. Y. t1. t2, P2」を代表アドレス・代表ポート番号とすることでもよい。

【0020】各サブフローの接続は、TCPの接続手順と同様である。複数のサブフローを立ち上げて束ねた

り、不要になったサブフローを解放したりする手順として、SETUP, ATTACH, DETACH, SHUTDOWN, RECOVERYの5つがある。

【0021】SETUPは、マルチリンクの接続を立ち上げる処理で、サブフローの接続を一つずつふやすsequential setupと、複数のサブフローの接続を並行に張って後で束ねていくconcurrent setupがある。

【0022】Sequential Setupでは、まず、一つ目の経路を通るサブフローについてMP (Multi-link Extension Protocol) _REQUESTのオプションを付けてTCPの3-ウェイ・ハンドシェイクによるセットアップを行う。通信相手がトランスポート層マルチリンクに対応している場合には、3-ウェイ・ハンドシェイクの際にMP_GRANTEDのオプションを付けたACKを返す。このMP_GRANTEDのメッセージには、トータルフローを識別する識別子（フロー識別子）が付与されている（接続の各方向につき一つずつIDが付与される）。一つ目の経路への接続に失敗した場合は、他の経路を通るサブフローを選んで同様のセットアップを行う。Sequential Setupの場合は、この最初の接続のローカルアドレス・ポートが代表アドレス・代表ポートとなる。サブフローのセットが完了したら、そのサブフローを使って通信を開始する（各パケットには通常のシーケンス番号に加えトータルシーケンス番号が付けられる）。他の経路のサブフローに関しては、後で述べるATTACH処理によりサブフローを追加していく。図7にSequential Setupの概念図を示す。

【0023】一方、Concurrent Setupでは、複数ある経路のうちの1より大きい任意の本数のサブフローの接続を同時にセットアップする。この場合、まず、同時にセットアップする各経路に関して、ローカル側のアドレスとポート番号を取得する。そして、各サブフローのセットアップの際に、MP_REQUESTのオプションにフロー識別子を付ける。このフロー識別子は、複数のTCP接続（複数のサブフロー）が一つのマルチリンク接続であることを識別ために利用される。接続を受ける方がトランスポート層マルチリンクに対応している場合には、最初に届いた接続要求の相手側アドレスとポートをその接続の相手側（接続開始側）の代表アドレス・代表ポートとし、MP_GRANTEDのオプションにその代表アドレス・代表ポートとフロー識別子を付けて3-ウェイ・ハンドシェイクのACKを返す。二つ目以降のサブフローについては、トータルフローの代表アドレス・代表ポートとフロー識別IDを付けたMP_GRANTEDをACKと共に返す。図8にConcurrent Setupの概念図を示す。

【0024】ATTACHではサブフローの追加を行う（図7参照）。ATTACHの際には、マルチリンクに対応している相手に対してMP_ATTACHのオプションとフロー識別子を付けてサブフローの接続を行う。

相手側は代表アドレス・代表ポートで表されるトータルフローがまだオープンな状態ならば、MP_ATTACHEDを付けてACKを返す。サブフローのセットアップが完了したら、トータルフローの一部のデータを開始する（以後、各パケットには通常のシーケンス番号に加えトータルシーケンス番号が付く）。このATTACHによるサブフローの追加は、前述のSequential Setup以外にも、トータルフローの転送量が突然増えた時あるいは他のサブフローの状態が大きく変動した時や、一定時間での周期的なトリガ等で発生することが考えられる。

【0025】DETACHはサブフローの削除を行う。DETACHによるサブフローの削除は、各サブフローについては通常のTCPの接続の終了と同じである。但し、代表ポート番号として使われているサブフローをDETACHする際には、トータルフローが終了するまで、そのポート番号は再利用出来ない。サブフローの削除は、サブフローを追加してもトータルの性能が変わらない場合や、転送が他のフローに比して小さい場合などに発生する。

【0026】SHUTDOWNはトータルフローでの通信を完了する処理である。SHUTDOWNの際には、各サブフローを通常のTCPと同様に終了するとともに、サブフローは既に終了しているのに代表ポート番号として当該ポート番号を使っている場合には、そのポート番号を解放する。

【0027】RECOVERYは、突然、サブフローが切断された場合に行われる処理である。定常状態では各サブフローで再送処理を行っているので、トータルフローではトータルシーケンス番号の順番にデータを並べるだけで良いが、通信途中で突然サブフローが切断された場合だけは、RECOVERY処理として、届いたことが確認されていない時点以降のデータをトータルフローとして再送処理する。

【0028】最後に、以上の処理を実現する為に必要されるネットワーク層より下位層について説明する。

【0029】以上に述べたような本発明のトランスポート層マルチリンク通信を実現するためには、各外部接続毎に複数のネットワーク層のネットワークを構成する必要がある。この解決には「物理データリンクの多重化」「論理データリンクの多重化」「IPネットワークの多重化」といった手段が考えられる。「物理データリンクの多重化」は、図9に示すように、各々の外部接続毎に一つずつの物理データリンクを用意し、エンドホストがそれぞれのデータリンクに複数接続するものである。例えば、一つのホスト10が複数のイーサネット（登録商標）インターフェイスを持ち、それぞれを異なる物理ネットワークに接続するような形態である。「論理データリンクの多重化」は、図10に示すように、データリンク層の用意する仮想LAN（Virtual LAN, VLA

N）機能を利用して、単一の物理インターフェイスを複数の仮想的な論理データリンクに分割して、各々の論理データリンク毎に各外部接続に繋がるネットワークとして構築するものである。具体的には、IEEE802.1Qに見られるEthernet（登録商標）VLANやATMのVC（Virtual Circuit）がこれに相当する。「IPネットワークの多重化」は、図11に示すように、各ホストに各外部接続毎の複数のIPアドレスを付与し、中間のルータが発信元のIPアドレスと送信先のIPアドレスの組みをもって経路制御するものである。このようなルータ機能として具体的には、routeやSuMiRe等が挙げられる。

【0030】以上、本発明を実施形態に基づき具体的に説明したが、本発明は、かかる実施形態に限定されるものでなく、その要旨を逸脱しない範囲において種々変更可能であることは勿論である。例えば、実施形態ではエンドホスト間の通信を仮定したが、ルータが物理的もしくは論理的に複数のアドレスを持っている場合、ルータ間においても同様にマルチリンク通信が可能である。また、実施形態ではTCPを拡張してマルチリンク接続を実現する例を説明したが、UDPを拡張しても同様にマルチリンク接続を実現できる。

【0031】

【発明の効果】以上のように、本発明によれば、複数の外部接続を持つネットワークから外部ネットワークへの通信において、複数の外部接続経路毎の通信をトランスポート層で束ねることにより、複数の外部接続の効率的利用や、実際のデータ転送の状況に応じた動的な通信経路の選択が可能になる。これにより、インターネットなどにおける通信をより効率良くまた高速に行うことができる。

【図面の簡単な説明】

【図1】本発明のトランスポート層マルチリンク通信のプロトコルスタックの概念図である。

【図2】通常のTCP/IPスイートにおけるプロトコルスタックを示す図である。

【図3】本発明のトランスポート層マルチリンク通信のアドレスとパケット転送の関係を示す図である。

【図4】本発明のトランスポート層マルチリンク通信のサブフローとトータルフローの関係を示す図である。

【図5】本発明のトランスポート層マルチリンク通信のサブフローシーケンス番号とトータルフローシーケンス番号の関係を示す図である。

【図6】図5のサブフローシーケンス番号とトータルシーケンス番号の関係を説明するためのデータ構成を示す図である。

【図7】本発明のトランスポート層マルチリンク通信におけるSequential Setupの概念図である。

【図8】本発明のトランスポート層マルチリンク通信におけるConcurrent Setupの概念図である。

【図9】本発明のトランスポート層マルチリンク通信を実現する物理データリンクの多重化構成を示す図である。

【図10】本発明のトランスポート層マルチリンク通信を実現する論理データリンクの多重化構成を示す図である。

【図11】本発明のトランスポート層マルチリンク通信を実現するIPネットワークの多重化構成を示す図である。

る。

【符号の説明】

10, 20 ホスト

30 ネットワーク

101, 102 ルータ

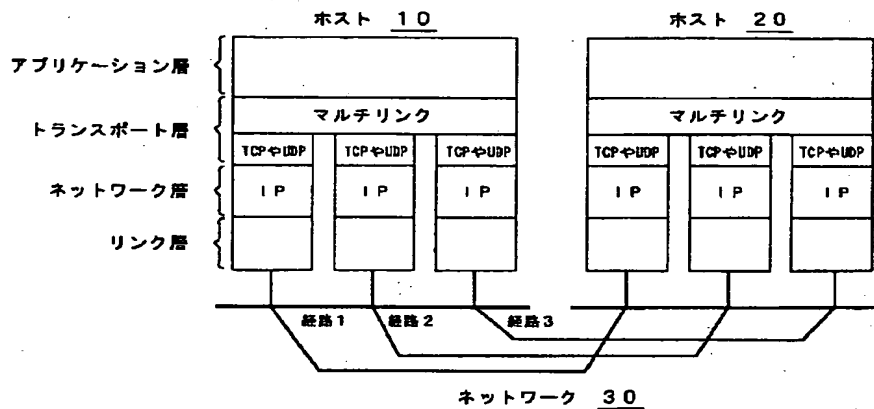
110, 120 パケット

200 トータルフロー

210, 220 サブフロー

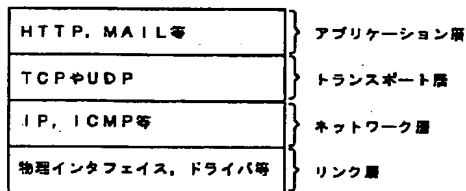
【図1】

本発明のトランスポート層マルチリンクの
プロトコルスタック（外部接続が3の場合）



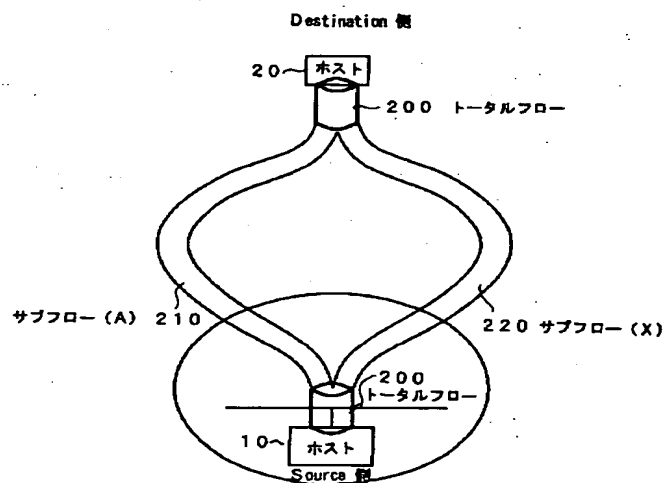
【図2】

従来のプロトコルスタック



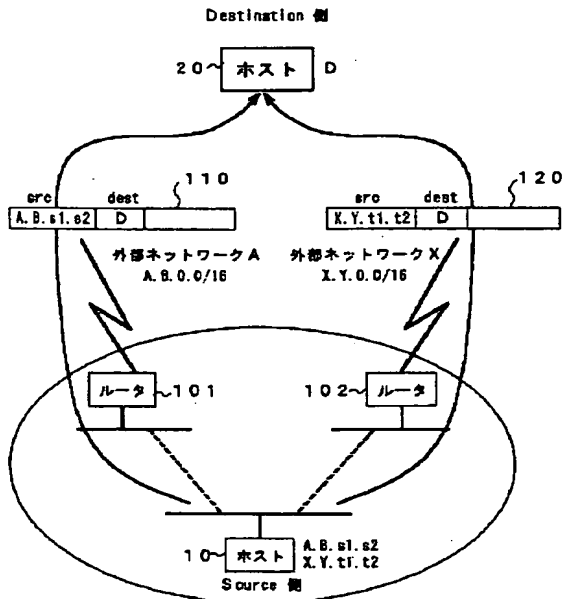
【図4】

サブフローとトータルフロー



【図3】

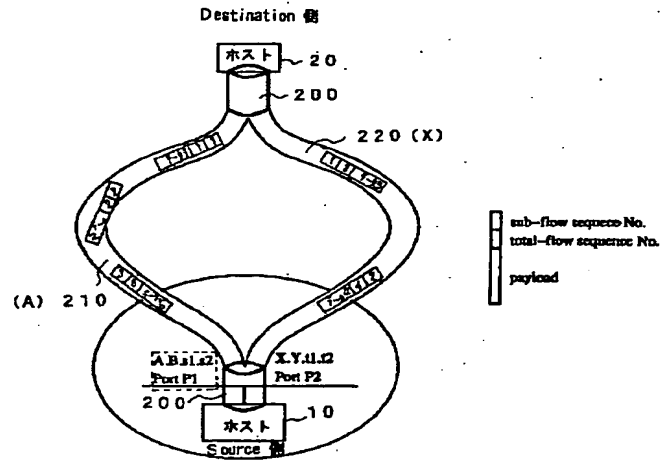
トランスポート層マルチリンク
のアドレスとパケット転送



(アドレスは一例、外部接続が2の場合)

【図5】

TCPマルチリンクのシーケンス番号



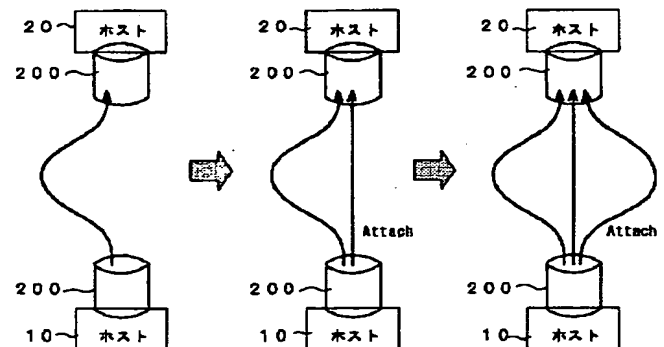
【図7】

【図6】

データのシーケンス番号

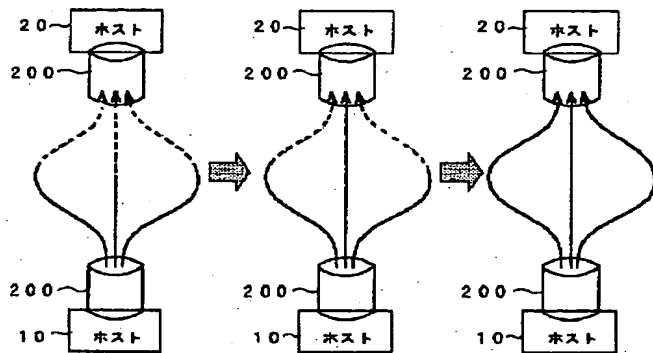
サブフロー シーケンスNo.	1 (A)	2 (A)	1 (X)	2 (X)	3 (A)
	データ1	データ2	データ3	データ4	データ5
トータル シーケンスNo.	1	2	3	4	5

Sequential Setup



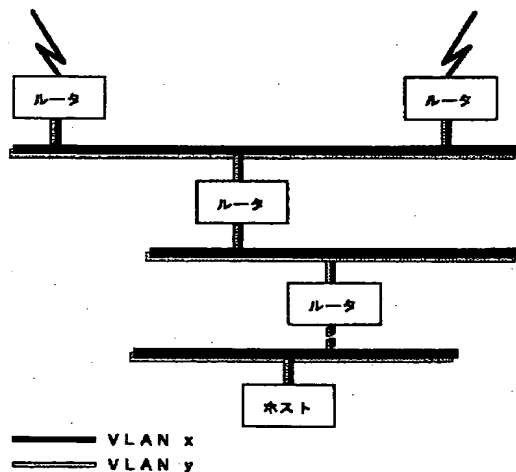
【図8】

Concurrent Setup



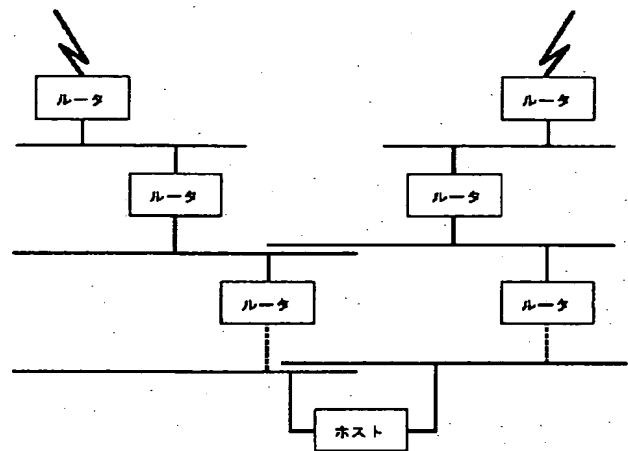
【図10】

論理データリンクの多重化



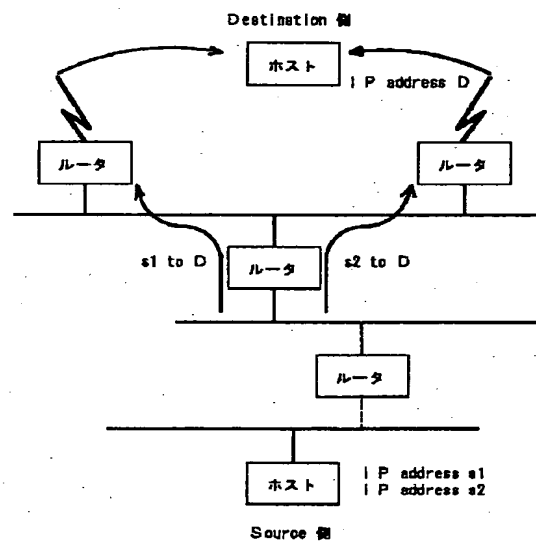
【図9】

物理データリンクの多重化



【図11】

I Pネットワークの多重化



フロントページの続き

(72)発明者 高田 敏弘
東京都千代田区大手町二丁目3番1号 日
本電信電話株式会社内
(72)発明者 福田 健介
東京都千代田区大手町二丁目3番1号 日
本電信電話株式会社内

Fターム(参考) 5K030 GA03 HC01 HC13 HD03 LB06
LE03 MB13
5K033 AA01 AA03 BA05 CC01 DA06
DB18
5K034 AA01 AA07 DD03 KK28
9A001 CC02 CC06 JJ12 JJ25 KK56